1978 ISSN 0017 – 4939

HERMANN SCHROEDEL VERLAG KG
Hannover · Dortmund · Darmstadt · Berlin

Gesamtherstellung: Druckerei Hans Oeding, Braunschweig

Printed in Germany

H 6661 F

Grundlagenstudien aus Kybernetik und Geisteswissenschaft

Erste deutschsprachige Zeitschrift für Kybernetische Pädagogik und Bildungstechnologie

Informations- und Zeichentheorie
Sprachkybernetik und Texttheorie
Informationspsychologie
Informationsästhetik
Modelltheorie
Organisationskybernetik
Kybernetikgeschichte
und Philosophie der Kybernetik

Begründet 1960 durch Max Bense
Gerhard Eichhorn
und Helmar Frank

Band 19 · Heft 3
September 1978
Kurztitel: GrKG 19/3

INHALT

KYBERNETISCHE FORSCHUNGSBERICHTE

HERMANN SCHROEDEL VERLAG KG

Im Verlaufe der sechziger Jahre gewann im deutschen Sprachraum, insbesondere im Umkreis der „Grundlagenstudien aus Kybernetik und Geisteswissenschaft", die Erkenntnis an Boden, daß die eigentliche Triebfeder der Kybernetik das Bedürfnis ist, die Vollbringung auch *geistiger* Arbeit an technische Objekte zu delegieren, kurz: sie zu *objektivieren*, und daß dies nicht ohne eine über die geisteswissenschaftlich-phänomenologische Reflexion hinausgehende wissenschaftliche Anstrengung in vorhersehbarer und reproduzierbarer Weise möglich ist, nämlich nicht ohne eine *Kalkülisierung* geistiger Arbeit. Die Bedeutung der Logistik, der Informationstheorie und der Theorie abstrakter Automaten als mathematische Werkzeuge wird von diesem Gesichtspunkt aus ebenso einsichtig wie der breite Raum, den die Bemühungen um eine Kalkülisierung im Bereich der *Psychologie* und im Bereich der Sprache bzw., allgemeiner, der *Zeichen*, einnahmen.

Die geistige Arbeit, deren Objektivierbarkeit allmählich zum Leitmotiv dieser Zeitschrift wurde, ist nicht jene geistige Arbeit, die sich selbst schon in bewußten Kalkülen vollzieht und deren Objektivierung zu den Anliegen jenes Zweiges der Kybernetik gehört, die heute als Rechnerkunde oder Informatik bezeichnet wird. Vielmehr geht es in dieser Zeitschrift vorrangig darum, die verborgenen Algorithmen hinter jenen geistigen Arbeitsvollzügen aufzudecken oder wenigstens durch eine Folge einfacherer Algorithmen anzunähern und damit immer besser objektivierbar zu machen, welche zur Thematik der bisherigen Geisteswissenschaften gehören. Der größte Bedarf an Objektivation in diesem Bereiche ist inzwischen bei der geistigen Arbeit des *Lehrens* aufgetreten. Mit der Lehrobjektivation stellt diese Zeitschrift ein Problem in den Mittelpunkt, dessen immer bessere Lösung nicht ohne Fortschritte auch bei der Objektivierung im Bereich der Sprachverarbeitung, des Wahrnehmens, Lernens und Problemlösens, der Erzeugung ästhetischer Information und des Organisierens möglich ist. Die Bildungstechnologie als gemeinsamer, sinngebender Bezugspunkt soll künftig auch bei kybernetikgeschichtlichen und philosophischen Beiträgen zu dieser Zeitschrift deutlicher sichtbar werden. (GrKG 13/1, S. 1 f.)

Manuskriptsendungen gemäß unseren Richtlinien auf der dritten Umschlagseite an die Schriftleitung:

Prof. Dr. Helmar Frank
Assessorin Brigitte Frank-Böhringer
(Geschäftsführende Schriftleiterin)
Institut für Kybernetik
Heiersmauer 71, D-4790 Paderborn
Tel.: (0 52 51) 3 20 23 u. 2 14 56

Die GrKG erscheinen in der Regel mit einer Knapptextbeilage in Internationaler Sprache mit dem Titel „Homo kaj Informo".

Anzeigenverwaltung und Vertrieb: Hermann Schroedel Verlag KG, Zeißstraße 10, D-3000 Hannover 81

Erscheinungsweise: Viermal im Jahr mit je ca. 36 Seiten.
Preis: Einzelheft DM 9,–, Jahresabonnement DM 32,–.
Jeweils zuzüglich Versandspesen. Alle Preise enthalten 6 % MWSt.

Abbestellungen von Jahresabonnements nur bis einen Monat vor Jahresende.

Netzwerktopologie von Informationsverbundsystemen

von Christel NIEDEREICHHOLZ, Bad Homburg
aus der Gesellschaft für Angewandte Informatik (GAI)
(Geschäftsführerin: Dr. Christel Niedereichholz)

1. *Einleitung und Problemstellung*

Seit einigen Jahren sind die Grenzen der mittels Einzelrechensystemen möglichen Datenverarbeitung erkannt. Wissenschaftliche Institutionen, Hersteller und Anwender entwickeln Konzepte zum Verbund von Rechensystemen aus folgenden Motivationen heraus:

a) Datenverbund – Verschiedene, in einem Netz gekoppelte Rechner können gemeinsam Daten und Programme benutzen. Dateien und Datenbanken (distributed data bases) können von mehreren Installationen gemeinsam genutzt und gepflegt werden.
b) Kapazitätsausgleich – Bei Überlastung einzelner Systeme kann über das Verbundnetz ein Kapazitätsausgleich erfolgen.
c) Hardwarenutzung – In einem Verbundsystem kann ein Betriebsmittelausgleich in der Form erfolgen, daß bestimmte Terminals (Endgeräte) und Speicher zu gemeinsamer Nutzung zur Verfügung stehen.
d) Softwarenutzung – Spezielle Software, die nur bei einzelnen Netzteilnehmern implementiert ist, kann von allen gekoppelten Rechensystemen genutzt werden.
e) Kommunikationsverbund – Ein Verbundsystem ermöglicht neue Formen der Kommunikation, wie beispielsweise Telekonferenzen und computergestützte Ausbildung.

In der vorliegenden Arbeit werden Strukturen, die aufgrund der verschiedenen Typen der sogenannten „Netzwerkmaschine" klassifiziert werden, dargestellt. Die Netzwerkmaschine (Bild 1) läßt unterscheiden zwischen dem Verbundsteuersystem zur Lösung von Koordinations- und Transformationsaufgaben und dem Datenübertragungssystem, das im Netzwerk die Vermittlungs- und Übertragungsfunktion erfüllt.

Ein Informationsverbundsystem stellt ein Netzwerk dar, in dem verschiedene Rechnersysteme als Knoten durch Kommunikationswege miteinander verbunden sind. In den Rechnersystemen laufen Prozesse ab, die entweder einen Informationsfluß einleiten (Quellprozesse), oder zur Abgabe von Informationen aufgefordert werden (Zielprozesse).

Die möglichen Strukturen und Abläufe in einem Informationsverbundsystem können durch die sogenannte Netzwerkmaschine (Howe, Kibler, 1971) dargestellt werden. Die Funktionen dieser Modellmaschine werden realiter durch Kommunikations- oder

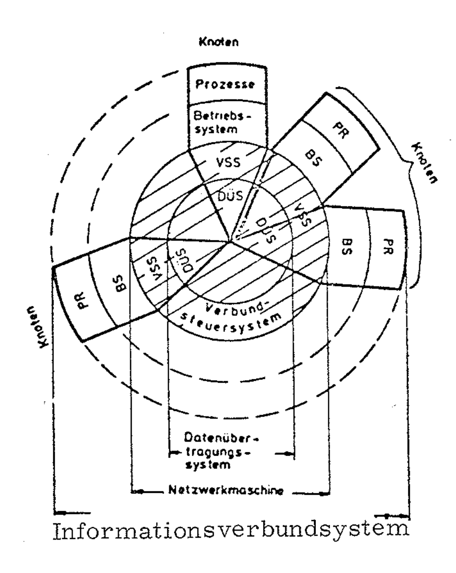


Bild 1: Informationsverbundsystem und Netzwerkmaschine

Verbundrechner erfüllt, wie zum Beispiel durch den Interface Message Processor (IMP) im Advanced Research Projekt Agency (ARPA)-Netzwerk.

Aus Bild 1 geht hervor, daß ein Informationsverbundsystem aus der Netzwerkmaschine und mehreren Knoten besteht. Ein oder mehrere Rechner werden dadurch zu Knoten des Informationsverbundsystems, daß ihnen ein Teil der Netzwerkmaschine zugeordnet wird. Die Funktionen der Netzwerkmaschine werden von zwei Teilbereichen, dem Verbundsteuersystem und dem Datenübertragungssystem erfüllt. Erfolgt die Zuordnung der Steuerungsfunktion gleichmäßig auf alle Knoten, so liegt eine dezentrale Steuerung vor. Das Netz wird zentral gesteuert, wenn nur ein Knoten die Verbindungen im System herstellt.

Im folgenden soll zunächst versucht werden, die vollständige Klassifikation der möglichen Verbundsteuersysteme, danach eine solche der Datenübertragunssysteme zu geben, um schließlich in Abschnitt 4 noch einen Ansatz zu einer Klassifikation der Realisierung der Netzwerkmaschinen-Strukturen zu machen.

2. *Verbundsteuersystem*

Das Verbundsteuersystem hat zwei Aufgaben:

a) Es werden logische Verbindungen zwischen Quell- und Zielprozeß hergestellt und abgebaut und der Informationsfluß überwacht (Koordination).
b) Es werden Aufträge an das Zielsystem und die Daten für den Zielprozeß in eine für das Rechensystem passende Form gebracht (Transformation).

Von besonderem Interesse sind die äußeren und inneren Schnittstellen des Verbundsteuersystems. An den äußeren Schnittstellen mit den Betriebssystemen findet der Nachrichtenaustausch mit den Prozessen statt. Die Informationen aus den Quellprozessen müssen in Warteschlangen verwaltet werden, bevor sie an das Datenübertragungssystem weitergeleitet werden. Die von den Zielprozessen rückübermittelten Nachrichtenbestätigungen dienen gleichzeitig der Aufdeckung von Folgefehlern. An den inneren Schnittstellen mit dem Datenübertragungssystem organisiert das Verbundsteuersystem die Informationsübermittlung durch Festlegen von Nachrichtenlängen und Aufteilung einzelner Nachrichten (package switching).

Da es noch keine einheitliche Steuersprache für die in Informationsverbundsystemen gekoppelten Rechner gibt, hat das Verbundsteuersystem in überwiegendem Maße Transformationsaufgaben zu erfüllen. Dabei sind sowohl die Aufträge vom Quell- an den Zielprozeß, als auch die Daten zu transformieren. Diese umfangreichen Transformationsaufgaben des Steuersystems führen zu der Forderung nach einer einheitlichen Normsteuersprache, mit der sich Aufträge an die Netzwerkmaschine und die Knoten sowie Anweisungen zur Datentransformation formulieren lassen.

3. *Datenübertragungssystem*

Den Kern der Netzwerkmaschine stellt das Datenübertragungssystem dar, in dem einige Determinanten herstellermäßig (IBM–SDLC) oder durch Standards (CCITT–X.25) festliegen.

Die Determinanten sind:
a) die Netztopologie,
b) das Übertragungskonzept und
c) die vorgegebenen Leitungsprozeduren.

Die Topologie eines Netzes ist gegeben durch die geographische Lage der Knoten und ihrer Verbindungen. Es existieren zwei Grundstrukturen, das Sternnetz (Bild 2) und das Maschennetz (Bild 3), sowie mehrere aus den beiden Grundstrukturen kombinierte Formen (Bild 4).

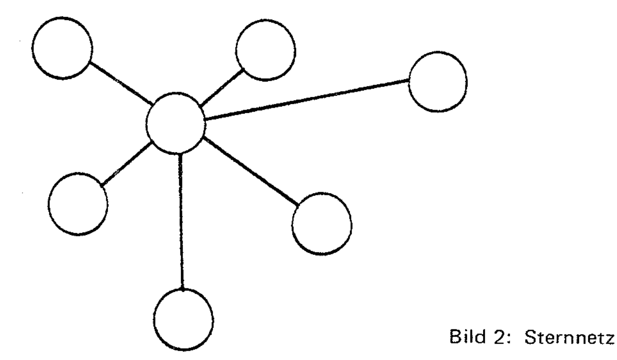


Bild 2: Sternnetz

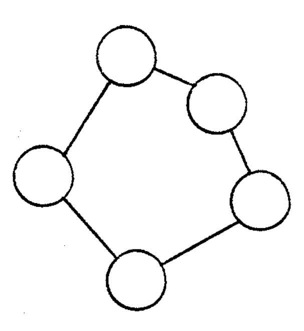
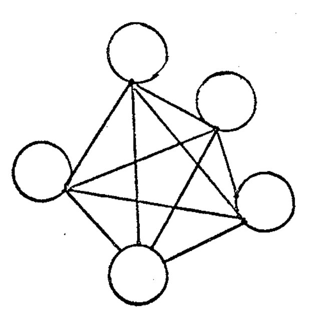



Bild 3: Maschennetze

In einem Sternnetz ist jeder Knoten mit der Zentrale verbunden. Den Vorteilen der einfachen Verwaltung und des kurzen Wegenetzes stehen die Nachteile des hohen Ausfallrisikos der Zentrale gegenüber. Diesen Nachteil kann man durch erhöhten Kostenaufwand überwinden, indem die Zentrale als Mehrprozessorsystem realisiert wird.

In einem Maschennetz ist jeder Knoten mit mindestens zwei anderen verbunden. Im einfachsten Fall einer solchen Verbindung spricht man von einem Ringnetz. Bestehen in einem Netzwerk mit N Knoten für jeden Knoten $N-1$ Verbindungen zu allen übrigen Knoten, so spricht man von einem vollvermaschten Netz. Zwischen diesen beiden Extremen ist jeder Grad der Vermaschung möglich. Der Vorteil der Maschennetze

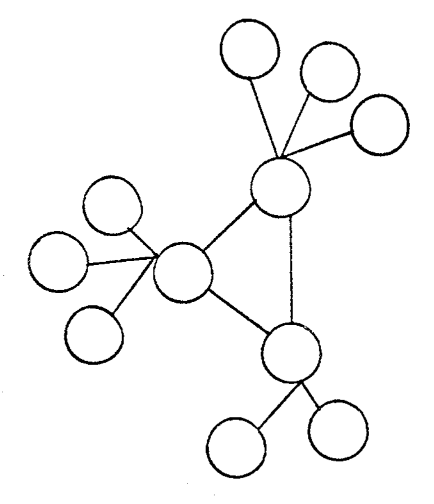
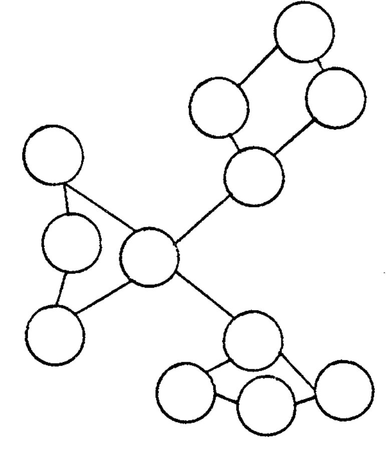



Bild 4: Kombinationen der Grundstrukturen

gegenüber den Sternnetzen besteht darin, daß jeder Knoten mit mindestens zwei Leitungen an andere Knoten gekoppelt ist, wodurch das Ausfallrisiko des Gesamtnetzes im Vergleich zu den Sternnetzen wesentlich verrringert ist. Die Nachteile wachsen mit dem Grad der Vermaschung durch ansteigende Leitungskosten.

In den Kombinationen der beiden Grundstrukturen wird versucht, die Vorteile von Stern- und Maschentopologie zu verbinden und die Nachteile beider Netzarten zu verringern.

Die Leistungsfähigkeit des Datenübertragungssystems wird neben der Netztopologie durch das Übertragungskonzept bestimmt. Es sind zwei Organisationsformen von Übertragungswegen möglich, die Punkt-zu-Punkt-Verbindung (Bild 5) und die Mehrpunktverbindung (Bild 6).

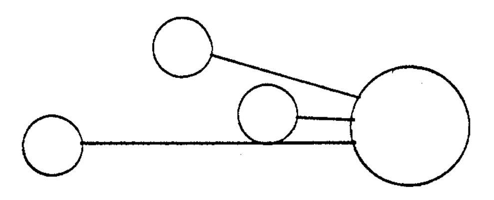


Bild 5: Punkt-zu-Punkt-Verbindung des Übertragungsweges

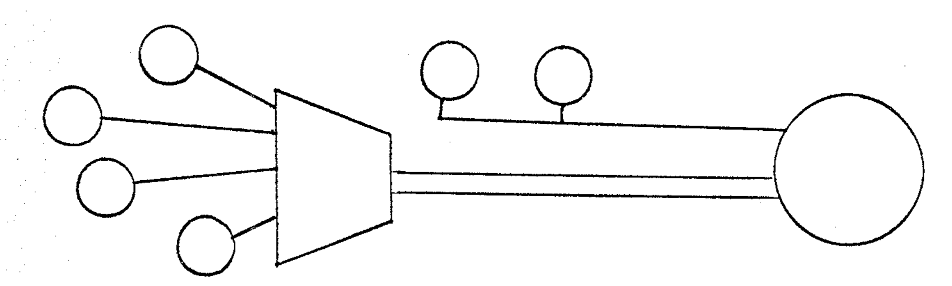


Bild 6: Mehrpunktverbindung des Übertragungsweges

Das Trapez symbolisiert einen Mehrpunktverbindungskonzentrator zwischen mehreren Endgeräten und Zentrale.

Das Übertragungskonzept besteht neben der Auswahl der Übertragungswege in der Auswahl der Leitungen. Der Übertragungsweg stellt die topologische Verbindung der Knoten dar, die Leistung ist deren technische Auslegung.

Die Bestimmungsgrößen bei der Auswahl der Übertragungsleitung sind:

a) Übertragungsdauer, die bestimmt wird durch das Daten-Volumen, die Übertragungsgeschwindigkeit und die Betriebszeit,

b) Übertragungskosten, die bestimmt werden durch Übertragungsdauer, Kosten der Datenübertragungseinheit, Adernzahl der Leitung, Leitungslänge, Leitungsart und Übertragungszeitpunkt,
c) Übertragungssicherheit, die bestimmt wird durch Fehlersicherungsverfahren, Störeinflüsse, Blocklänge, Übertragungsgeschwindigkeit, Modulationsart, Code und Redundanz,
d) Datenvolumen, das bestimmt wird durch Nachrichten pro Endgerät, Frequenz des Datenanfalls, Endgeräte pro Leitung, fehlerbedingte Wiederholungen, Redundanz, Ping-Pong-Effekt (fehlerhafter, wechselseitiger Rückübertragungs-Effekt, d.h. die Nachricht wird wie ein Ping-Pong-Ball vom Ziel zur Quelle zurück und wieder umgekehrt hin- und hergeleitet) bei Blockübertragung,
e) Antwortzeitverhalten, das bestimmt wird durch Anzahl der Endgeräte, Aufbauzeit der logischen Verbindungen, Betriebsweise und Datenvolumen.

Die Schaltung durch eine Leitung kann direkt, durchgeschaltet, wie zum Beispiel im Telefonverkehr sein, oder es besteht zwischen Quell- und Zielprozeß eine Zwischenschaltung von speichernden Vermittlungsstellen. Für Informationsverbundsysteme wird die letztere Schaltungsart realisiert. Eine Nachricht wird nicht als geschlossenes Ganzes, sondern in Pakete unterteilt übertragen. Jedes Paket erhält einen Vorspann (header), der in dem Vermittlungsknoten interpretiert wird, und eine gerichtete Weitervermittlung bis zum Zielknoten ermöglicht (Bild 7).

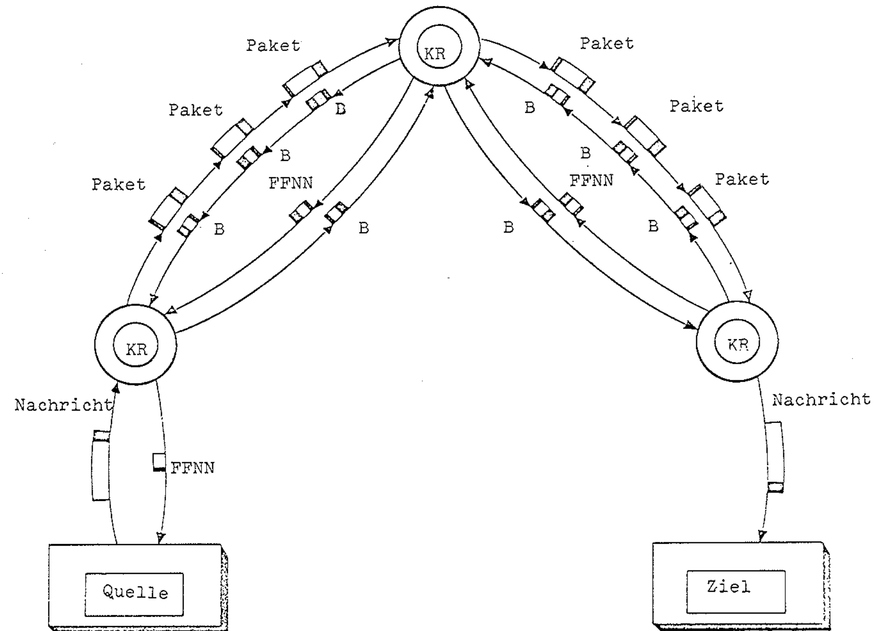


Bild 7: Übertragung der Information über Kommunikationsrechner (KR) im ARPA-Netz
B = Bestätigung, FFNN = Fertig-für-nächste-Nachricht, KR = Kommunikationsrechner

Der genaue Weg eines Pakets durch das Netzwerk kann vorbestimmt sein oder je nach Netzzustand von Knoten zu Knoten neu festgelegt werden. Diese Routenvorgabe von Knoten zu Knoten erfolgt durch interne Steuerbefehle, durch die jeder Knoten von der Auslastung seiner Nachbarn unterrichtet wird. Eine Problematik der Paketübertragung beruht darin, daß die Pakete nicht mehr in der ursprünglichen logischen Reihenfolge am Zielknoten ankommen. Diese Problematik wird dadurch gelöst, daß jeder Datenblock, wie schon in Bild 7 dargestellt, in seinem Kopfteil (header) eine entsprechende Numerierung erhält, die das logische Verknüpfen im Zielknoten ermöglicht (Bild 8).

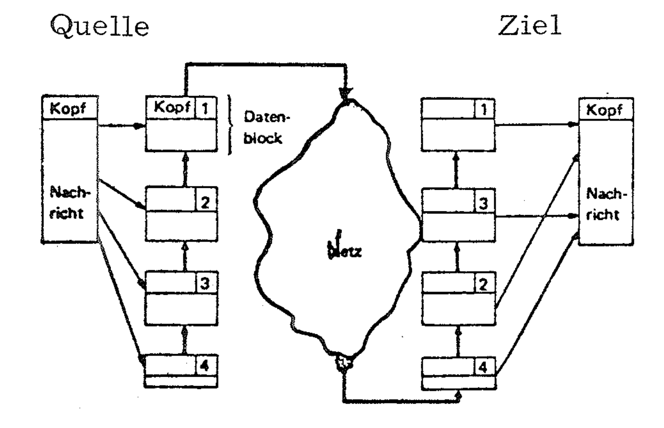


Bild 8: Logische Verknüpfung der Datenpakete im Zielknoten

4. *Modelltypen*

Die existierenden Rechnerverbundsysteme kann man nach der Art der Verbindung in verschiedene Modelltypen einteilen. Die einfachste Auslegung ist die direkte Kopplung über die hard- und softwaremäßigen Einrichtungen der Datenfernübertragung (DFÜ). Diese Verbundart ist von weniger großem Interesse, da nur gleichartige Rechner gekoppelt werden können (homogenes Netz, Bild 9). Mit steigender Knotenzahl

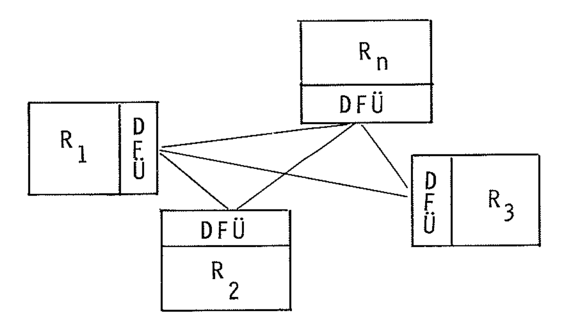


Bild 9: Einfache Rechnerkopplung durch Datenfernübertragung

wächst der Steuerungsaufwand für jeden der so verbundenen Rechner überproportional an. Erweiterungen und Verkleinerungen eines Netzes mit direkter Rechnerverbindung verursachen unverhältnismäßig hohe Kosten. Ein Beispiel hierfür ist die ASP-Mehrrechner-Koppelung bei IBM (ASP = Attached Support Processor).

Einen anderen Modelltyp stellt die Verbindung der einzelnen Rechensysteme über einen Vermittlungsrechner dar (Bild 10). Die Funktionen des Verbundsteuer- und Datenübertragungssystems sind aus den Rechensystemen ausgegliedert und hardwaremäßig dem Vermittlungs- oder Kommunikationsrechner zugeordnet, der damit die Aufgaben der Netzwerkmaschine übernimmt. Da diese Auslegung nur bei einer geringen Anzahl von teilnehmenden Rechensystemen möglich ist, können Anforderungen wie

- hohe Übertragungsleistung,
- kurze Antwortzeiten,
- hohe Flexibilität des Systems in bezug auf die verschiedenen Fabrikate der angeschlossenen Rechner,
- Simultanbetrieb von Ein- und Ausgabeoperationen,
- Datensicherung und Fehlererkennung,
- geringe System- und Systemveränderungskosten

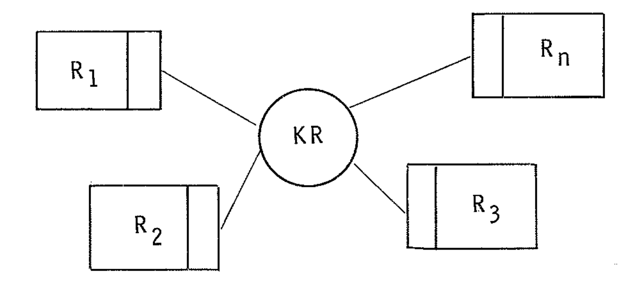


Bild 10: Verbundsystem mit einem zentralen Vermittlungsrechner (Kommunikationsrechner – KR)

erfüllt werden. Die Nachteile dieses Verbundsystems über einen zentralen Vermittlungsrechner liegen in der engen räumlichen Anordnung und der dadurch mitbedingten geringen Anzahl möglicher Rechner.

Eine höhere Flexibilität des Gesamtverbundes und eine größere geographische Ausweitung läßt sich durch die Zwischenschaltung je eines Vermittlungsrechners vor jedes Rechnersystem erzielen. Die Vermittlungsrechner können dabei ein autonomes Netz bilden, das heißt sie sind hardwaremäßig gleichartig und nivellieren dadurch technische Unterschiede der Verbundrechner. Die inneren Schnittstellen sind standardisiert, wodurch topologische Veränderungen des Netzes erleichtert werden. An einen Vermittlungsrechner können viele Teilnehmer angeschlossen werden (Bild 11). Das bedeutet,

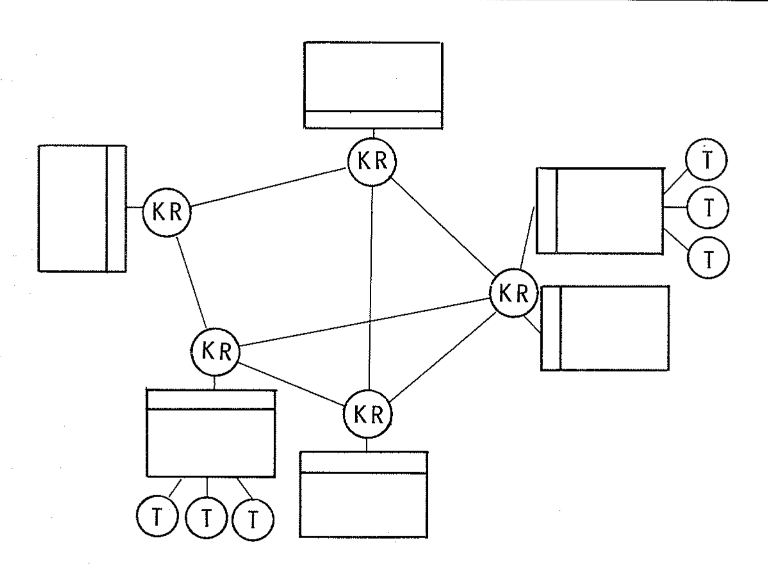


Bild 11: Informationsverbundsystem

daß die äußere Schnittstelle vom Vermittlungsrechner zu den Rechensystemen anpassungsfähig sein muß. Dialogfähige Terminals können sowohl an ein teilnehmendes Rechensystem als auch direkt an den Vermittlungsrechner angeschlossen werden, wenn dieser die notwendigen Steuerfunktionen übernehmen kann. Letzteres erlauben die sogenannten TIP-Rechner des ARPA-Netzes (TIP = Transaction Interface Processor), im Gegensatz zu den IMP-Netzrechnern, die eine Host-Anlage anschließen (vgl. hierzu Mahmoud, Riordon, 1976).

Schrifttum

Howe, W. G., Kibler, T. A.: Control concepts of a logical network machine, in: Computer Networks, Infotech Information 1971, S. 527 – 541

Mahmoud, S., Riordon, J. S.: Optimal allocation of Resources in distributed information networks, ACM Transactions on Database Systems, Vol. 1, No. 1, März 1976, S. 66 – 78

Eingegangen am 19. Januar 1978

Anschrift der Verfasserin:
Dr. Christel Niedereichholz, Frankenstr. 2, D-6380 Bad Homburg

Grundlagen und sprachpädagogische Anwendung einer informationstheoretischen Transferanalyse

von Helmar FRANK, Paderborn

aus dem Institut für Kybernetische Pädagogik des Landesforschungsinstituts FEoLL, Paderborn (Direktor: Prof. Dr. H. Frank)

1. *Problemstellung*

Die Informationspsychologie erforscht, und die kybernetische Pädagogik nutzt den Zusammenhang zwischen dem Informationsgehalt von Wahrnehmungsgegenständen bzw. Lehrstoffen einerseits und ihrer Aufnahmegeschwindigkeit in das Bewußtsein bzw. das Gedächtnis andererseits (vgl. die ausführlichen Darstellungen hierzu bei Frank, 1969.0.1; Kapitel 5; Weltner, 1970; von Cube 1970). Der Zusammenhang zwischen Lernzeit t und Lehrstoffinformation I ist u.U. sogar linear, z.B. in der neuerdings zunehmend benutzten „Lernzeitformel"

$$t = \frac{I}{\eta \cdot C_v} \ln \frac{1-p_o}{1-!p} \quad (1)$$

welche bei vorgegebener prozentualer Lehrstoffvorkenntnis p_o, bekannter Lerngeschwindigkeit C_v (in bit/sec) und bekannter Effikanz η ($o < \eta < 1$) eines rückkopplungsfreien oder nur schwach rückgekoppelten (z.B. eines Klassen-)Unterrichts die einzuplanende Lernzeit t liefert, nach welcher eine geforderte prozentuale Lehrstoffkenntnis $!p$ erwartet werden darf (Frank, 1975.A.5; 1976.A.1; 1976.H.1; 1976.H.2; 1977.A.1; 1977.A.5; Hilgers, 1978).

Der Informationsgehalt $I(z)$ einer Nachricht z wird im Mittel verringert, wenn eine andere Nachricht y, die von z nicht stochastisch unabhängig ist, vorher bekannt wird. Diese Abnahme heißt Transinformation. Aus ihrer Definition für den Fall zweier Einzelnachrichten y, z

$$t(yz) =_{Df} I(z) - I(z|y) = \mathrm{ld}\,\frac{1}{p(z)} - \mathrm{ld}\,\frac{1}{p(z|y)} \quad (2a)$$

$$= \mathrm{ld}\,\frac{1}{p(z)} - \mathrm{ld}\,\frac{p(y)}{p(yz)} = \mathrm{ld}\,\frac{p(yz)}{p(y)\cdot p(z)} = \mathrm{ld}\,\frac{p(zy)}{p(z)\cdot p(y)} = t(zy)$$

bzw. im Mittel für die Felder Y, Z, aus denen diese Nachrichten stammen:

$$T(YZ) =_{Df} H(Z) - H(Z|Y) = H(Z) - (H(YZ) - H(Y)) \quad (2b)$$

$$= H(Y) + H(Z) - H(YZ) = T(ZY)$$

folgt übereinstimmend die Symmetrie der Transinformation. Demnach gewinnt man im Mittel wie im Einzelfall aus einer Nachricht „über" eine andere Nachricht ebensoviel „Information" wie umgekehrt. Diese „Transinformation" (= „Information über ...") kann zwar im Einzelfall, nicht aber im Mittel negativ sein. (Ausführliche, elementare mathematische Begründungen finden sich z.B. in Frank, 1969.0.1, Kap. 3.1.)

Eine vorschnelle pädagogische Anwendung der Beziehung (1) zusammen mit (2a, b) könnte in der Begründung der These bestehen, daß das Erlernen eines Lehrstoffelements y die spätere Erlernung eines Lehrstoffelements z eines anderen Gebietes erleichtern oder erschweren könne, im Mittel jedoch allenfalls erleichtern, wobei die *Reihenfolge* der Erlernung der beiden Elemente oder Lehrgebiete *unerheblich* sei. Dies würde sämtliche Reihenfolgeprobleme der Lehrplanungstheorie aufheben und steht in striktem Gegensatz zu den Forschungsergebnissen über die zwischen verschiedenen Lehrstoffen bestehenden „natürlichen Ordnungen" (Lansky, 1971), die anschaulich in einem „Kohärenzstrukturdiagramm" (Weltner, 1975) zusammengefaßt sind und zu den wenigen unumstrittenen und grundlegenden Erkenntnissen der Didaktik und Lehrplanungstheorie gehören. Insbesondere wären sämtliche Lehrstoff*modelle* Lm sinnlos, da ja ihre Verwendung dadurch gerechtfertigt wird (Frank, 1974.A.3 und 1975.B.1), daß ihre Erlernung weniger Zeit Z kostet, als dadurch bei der *anschließenden* Erlernung des *eigentlichen* Lehrgegenstands Lg eingespart wird:

$$Z(Lm) + Z(Lg/Lm) < Z(Lg) \tag{3}$$

In der vorliegenden Arbeit bemühen wir uns (1) diesen Widerspruch aufzulösen, (2) die Lernerleichterung beim Aneignen eines Lehrstoffs Lg durch vorheriges Erlernen eines anderen Lehrstoffs – insbesondere eines Modells Lm von Lg – durch ein vom Begriff der Transinformation verschiedenes, jedoch informationspsychologisch fundiertes, nicht symmetrisches Transfermaß zu erfassen, und (3) die empirische Anwendbarkeit dieser informationstheoretischen Transferanalyse und ihre Tauglichkeit für ein bisher nicht befriedigend in der kybernetischen Pädagogik verankertes Teilgebiet der prospektiven Bildungswissenschaft nachzuweisen.

2. *Voraussetzungen informationstheoretischer Sätze und psychologisch-bildungspraktische Vorgaben*

2.1 *Endliches Gedächtnis oder wenigstens Ergodizität der Nachrichtenquellen*

Mehrere Aussagen der Informationstheorie setzen voraus, daß die zu betrachtenden Nachrichten als Folgen von Zeichen anzusehen sind, welche eine in ihren statistischen Eigenschaften unveränderliche Quelle mit „endlichem Gedächtnis" liefert, d.h., es wird angenommen, daß die Nachrichten Markoffketten sind, so daß es stets eine Zahl $m \geqslant 0$ gibt derart, daß die Wahrscheinlichkeit eines jeden von der Quelle gesendeten Zeichens höchstens von den m vorangegangenen Zeichen abhängt. Diese Voraussetzung erfüllt die natürliche Sprache weder auf der Stufe der Lautzeichen (Buchstaben) noch der Wortzeichen, da grundsätzlich in jedem Satz beliebig viele Einschiebungen möglich sind, ohne den Einfluß des Anfangs auf den Schluß des Satzes zu beseitigen. Mehrere informationstheoretische Theoreme konnten aber auch allgemeiner für alle „ergodischen" Quellen bewiesen werden; hier wird nach Definition die Abhängigkeit von einem früheren Zeichen wenigstens beliebig klein, wenn dieses nur genügend weit zurück liegt. Offensichtlich kann es eine solche Quelle weder in einer Welt geben, die einen Anfang in der Zeit hat (christliche Kosmogonie), noch in einer Welt der periodischen, ewigen Wiederkehr (buddhistische Kosmogonie). Dies schließt nicht aus, daß mit dem Modell der Markoffschen oder zumindest ergodischen Quelle dort gearbeitet werden kann, wo dieses Modell die Wirklichkeit hinreichend gut beschreibt; das kann vielfach auch im Bereich der Psychologie und der Pädagogik angenommen werden.

2.2 *Informationelle Akkomodation*

Im psychologischen-pädagogischen Bereich ist eine andere Voraussetzung der Informationstheorie meist nicht einmal annähernd erfüllt, nämlich die Annahme, der Empfänger sei mit den statistischen Eigenschaften der Quelle vertraut. Stattdessen erfolgt während des Wahrnehmens oder Lernens eine Annäherung der subjektiven Erwartungswahrscheinlichkeiten w des Empfängers an die vom Sender abhängigen Auftrittswahrscheinlichkeiten p der Zeichen. Daher ist im Mittel die subjektive Information zu groß und näherte sich erst allmählich an die gesendete Information an:

$$H_{\text{sub}} = \sum_i p_i \operatorname{ld} \frac{1}{w_i} \geqslant \sum_i p_i \operatorname{ld} \frac{1}{p_i} = H \tag{4}$$

Diese „informationelle Akkomodation" $H_{\text{sub}} \rightarrow H$ (Frank, 1960.A.3) verläuft zwar unbewußt, kann aber durch plötzlich entstehende Vermutungen oder Einsichten über die statistischen Eigenschaften der Quelle bzw. einer von ihr gesendeten Zeichenfolge (Nachricht) erheblich beeinflußt werden; die „kognitive Verlaufsfunktion" von Weltner (1977) veranschaulicht dieses. „Einsichten" über eine Nachricht z oder über den Nachrichtenfluß aus einer Quelle können im Empfänger offensichtlich durch eine andere Nachricht y geweckt werden, nämlich dann, wenn entweder

a) y als semantische Information just die relevanten Eigenschaften von z überträgt oder wenn
b) y die relevanten Eigenschaften von z aufweist, eine entsprechende informationelle Akkomodation bewirkt und damit den Empfänger in einen Zustand versetzt, in welchem er an die anschließend aufzunehmende Nachricht z schon (weitgehend) informationell akkomodiert ist.

2.3 *Beschleunigung der informationellen Akkomodation durch semantische Information*

Ein einfaches Beispiel für den Fall (a) läßt sich wie folgt konstruieren. Die Nachricht z bestehe aus einem deutschen Text der Länge 100 Wörter mit durchschnittlich 5 in

Majuskeln geschriebenen Buchstaben, wobei jedoch innerhalb jeden Wortes die Buchstabenfolge von hinten nach vorne geschrieben ist. Die subjektive Information dieses Textes dürfte anfangs bei 5 bit/Zeichen, nach Gewöhnung an die (nicht-bedingte) Häufigkeitsverteilung bei etwa 4 bit/Zeichen liegen, insgesamt also bei rund 2000 bit. Eine zweite Nachricht y bestehe nur aus dem Satz: „Die Nachricht z ist Wort für Wort von rückwärts zu lesen." Ein Ratetest nach Weltner (1970) könnte für diesen alleinstehenden Satz vielleicht $I(y) = 100$ bit Information ergeben. Die bedingte Information von z, wenn vorher y bekannt wurde, liegt (abhängig vom Inhalt) bei kaum mehr als 1 bit/Zeichen, also sicher unter 1000 bit. Damit gilt in diesem Beispiel:

(5a) $$I(y) + I(z|y) < I(z)$$

also nach (2a)

(5b) $$t(yz) = I(z) - I(z|y) > I(y).$$

Wegen der Proportionalität zwischen Wahrnehmungs- und Lernzeit einerseits und Information andererseits ist damit die Ungleichung (3) erfüllt: y ist eine gerechtfertigte Lernhilfe. Wird zuerst z wahrgenommen oder gelernt, dann dürfte dies auf die in y steckende subjektive Information keinen oder nur einen sehr geringen Einfluß haben; selbst wenn das Konstruktionsprinzip von z und damit die semantische Information von y schließlich erkannt wird, senkt dies die Information von y um weniger als 50 bit, keineswegs auf Null; sicher gilt also

(6) $$t(zy) = I(y) - I(y|z) < I(y)$$

Es ist also entgegen der Ableitung (2a) in diesem Falle

(7) $$t(yz) > t(zy).$$

Das Zutreffen von (5a) setzt zweierlei voraus:

1. die Nachricht z darf nicht so lang sein, daß die vor Abschluß der informationellen Akkomodation wegen (4) zuviel aufgenommene Information nicht ins Gewicht fällt; dann wird nämlich $I(z|y) \approx I(z)$;
2. die Nachricht y muß informationsarm sein, d.h. unter der großen Menge von eindeutig umkehrbaren Permutationsmöglichkeiten muß der Nachricht z eine solche zugrunde liegen (im Beispiel: die Reihenfolgeumkehrung), die subjektiv sehr wahrscheinlich ist, weil sie – ohne logisch-mathematische Notwendigkeit! – noch als „Ordnung" statt als „willkürlich" *empfunden* wird (Frank, 1971.A.2); die Beschreibung y irgendeiner „ungewöhnlichen", aber an sich gleichwertigen sonstigen Permutation z wäre subjektiv informationsreich, so daß in (5a) die linke Seite die rechte übertreffen könnte.

Der scheinbare Widerspruch zwischen (2a, b) und (3), der mit (7) offenkundig wird, läßt sich bei sorgfältiger Verfolgung der Umformungskette (2a) auflösen. Benutzt wird

dort die Symmetrieaussage

(8) $$p(yz) = p(zy).$$

In der Tat ist die Wahrscheinlichkeit dafür, daß im Feld Y die Nachricht y und im Feld Z die Nachricht z ausgewählt wird, gleich der Wahrscheinlichkeit, daß im Feld Z die Nachricht z und im Feld Y die Nachricht y ausgewählt wird, denn die UND-Verknüpfung von Aussagen ist kommutativ. Das bedeutet aber nicht, daß für die unbewußten subjektiven Wahrscheinlichkeiten, welche die jeweilige subjektive Information und damit die Wahrnehmungs-, Reaktions- und Lernzeit bestimmen, auch $w(yz) = w(zy)$ gelten müßte: die subjektive Wahrscheinlichkeit des Nachrichtenpaares kann davon abhängen, welche ihrer beiden Komponenten *zuerst aufgenommen* wird, während objektiv und damit in (8) eine solche Zeitfolge keine Rolle spielt! –

2.4 *Beschleunigung der informationellen Akkomodation durch Strukturerkennung*

In frühen Überlegungen zur Informationsästhetik (Frank, 1959.N.1, § 4.34–4.36) wurde bereits neben der Beschleunigung der informationellen Akkomodation durch semantische Information (Fall a von Abschnitt 2.2) auch die Beschleunigung durch Strukturerkennung (Fall b) angesprochen, wenngleich noch ohne saubere Unterscheidung zwischen den beiden Fällen und zwischen den Nachrichten y und z. Im Unterschied zu den Voraussetzungen der mathematischen Informationstheorie wird die Nachricht (z.B. Sprache oder Musik) nicht als Zeichenkette eines festen Repertoires (z.B. Laute oder Töne) betrachtet, sondern es wird unterstellt, daß der Empfänger durch einen „Birkhoffschen Übergang" („Superierung") auf eine Ebene von „Superzeichen" (z.B. Wörter oder Motive) die Nachricht auch als im allgemeinen wesentlich kürzere Kette im allgemeinen informationsreicherer Zeichen wahrnehmen kann. Dabei verringert sich insgesamt die Information, da (1) nicht alle statistisch möglichen Unterzeichenkombinationen zum Superzeichenrepertoire gehören (auf diesen Aspekt des „komplexbildenden Superierens" konzentrierte sich eine von F. von Cube, 1961.A.1, ausgelöste, jahrelange Diskussion; die Beiträge dazu gaben Lánský und Polák, 1977, als Sammelband heraus), und da (2) verschiedene solche Kombinationen (Aussprachemöglichkeiten desselben Wortes; Variationen eines Themas) dasselbe Superzeichen repräsentieren, so daß zwischen ihnen nicht mehr unterschieden wird („klassenbildendes Superieren"). Erst nach dem „Molesschen Übergang" wird die spezielle Realisierung des erkannten Superzeichens apperzipiert, wobei unterstellt wird, daß diese, als „ästhetische Information" bezeichnete, bedingte Information zusammen mit der Superzeicheninformation zumindest im Mittel wesentlich *kleiner* ist als die vor dem Birkhoffschen Übergang angefallene „Rohinformation". Dies widerspricht zwar scheinbar dem leicht beweisbaren allgemeinen Satz der Informationstheorie

(9) $$H(Y) + H(Z|Y) \; (= H(YZ)) \geqslant H(Z)$$

entspricht jedoch der Erfahrung, daß „alles ganz einfach wird", sobald man die strukturierenden Ordnungsprinzipien (die „Makrostruktur") erkannt und damit (von der

Ebene der Superzeichen aus:) einen „Überblick" gewonnen hat. Denn die „vom Menschen vermöge seiner Ordnungsgesichtspunkte vorgesehenen Möglichkeit bilden winzige Inseln innerhalb eines Ozeans an sich ebensogut möglich ... Kombinationen ... ‚Inseln' der vermeintlichen Ordnung im ‚Rauschen' des ‚Ozeans der Unordnung' ..." (Frank, 1969.0.1, § 2.1). Diese Veranschaulichung macht evident, daß (zumindest im Mittel) die Information darüber, auf welcher „Insel" der Standort (die spezielle Zeichenkombination) liegt – also $H(Y)$ – vermehrt um die Information über den Standort auf dieser Insel (bedingte Information) – also $H(Z|Y)$ – wesentlich *kleiner* ist als die Informationsangabe über irgendeinen Ort im „Gesamtbereich von Ozean mit Inseln" (d.h. über irgendeine mögliche Kombination der Unterzeichen). Das gilt aber nur, wenn (1) jede Insel (genauer: ihre Lage) und (2) der Umstand bekannt ist, daß der Standort nur auf einer dieser Inseln liegen kann – und wenn es sich um echte, d.h. durch den Ozean voneinander getrennte Inseln handelt. Dagegen vergrößert sich $H(Y)$ bzw. $H(Z|Y)$ so erheblich, daß (9) gilt, falls jeder Punkt möglicher Inselmittelpunkt sein kann, bzw. der relativ zu ihm beschreibbare Standort auch außerhalb der Insel liegen kann, bzw. die Inseln beliebig groß und damit überlappend sein können. Der Widerspruch zur Ungleichung (9) entsteht also dann, wenn man die linke Seite für einen Empfänger im Zustand *nach* vollzogenem Birkhoffschen Übergang berechnet, die rechte Seite dagegen im *vorangegangenen* Zustand, als noch nicht beachtet wurde, daß die im Ozean liegenden Punkte (also fast alle!) auszuschließen sind. Damit werden aber die mathematischen Voraussetzungen von (9) verletzt.

2.5 *Eigenschaften von Lehrstoffmodellen*

Zur Erklärung der Ungleichung (3) interpretieren wir das Lehrstoffmodell Lm als „Mittelpunkt" einer „Insel" z, d.h. als Repräsentant eines Superzeichens, der für einen Lerner durch wenig Informationsvermittlung zu beschreiben ist. (Vergleiche: Die Ortsangabe in einem cartesischen Korrdinatensystem ist subjektiv informationsarm, wenn es sich um einen Gitterpunkt handelt. In einem Musikstück wird ein Motiv in schlichter Form durch ein „Thema" eingeführt, dessen „Variationen" subjektiv informationsreicher sind.) Hierfür müssen Lehrstoffmodelle (über die Eigenschaften aller Modelle – vgl. Stachowiak, 1973 – hinausgehend) den Bedingungen der Einfachheit, Regelmäßigkeit und Deutlichkeit genügen (Frank, 1976.B.1; 1978).

Einfach ist Lm, wenn es den Blick auf das „Wesentliche" (die Struktur, das „Superzeichen") erleichtert statt davon durch informationsreiche Unwesentlichkeiten abzulenken. Als Modell einer Fläche nimmt man nicht einen Bilderbuchdeckel, dessen Dicke und Bebilderung in unwesentliche Dimensionen verweisen, sondern ein Polynom, eingezeichnet in ein cartesisches Koordinatensystem.

Regelmäßig ist Lm, wenn es Zufallsabweichungen von einem Grundmuster (Paradigma) vermeidet. Man wird also die Ecken des erwähnten Polynoms nicht abrunden und die Seiten nicht durch Wellenlinien ersetzen.

Deutlich ist Lm, wenn es die Schwierigkeiten getrennt darzubieten erlaubt. Man wird also als Polynom ein rechtwinkliges Dreieck wählen, dieses ausschneiden und zunächst mit den Katheten auf die Koordinatenachsen legen, um es später, nach erfolgter Flächenberechnung, nacheinander in Abszissen- und Ordinatenrichtung zu verschieben.

Die zu lernende Information $I(Lm)$ des Lehrstoffmodells ist unter diesen Voraussetzungen klein, und damit wegen (1) die Lernzeit $Z(Lm)$. Da die unmittelbare Quelle sämtlicher Lehrstoffe der Mensch ist, also eine Quelle, die mit dem Empfänger gleichartig ist, darf unterstellt werden, daß jeder Lehrgegenstand Lg eine „winzige Insel innerhalb des Ozeans" denkbarer Folgen von Lernelementen, also der Unterschied zwischen Lm und Lg relativ klein ist – und damit auch die noch zu lernende Information $I(Lg|Lm)$ über zusätzliche Besonderheiten, über Unregelmäßigkeiten und über wechselseitige Abhängigkeiten (Restriktionen) der Komponenten von Lg. Somit wird (3) – letzten Endes wegen der Ungleichung (4)! – ohne Widerspruch zu (1) und (2) erfüllbar.

Didaktisch kann die gemeinsame „Struktur" von Lm und Lg entweder durch Suchenlassen der Gemeinsamkeit von Lm und anderen, geistig leicht zugänglichen Repräsentanten derselben Struktur vermittelt werden (z.B. durch Konfrontation der Internacia Lingvo mit der Muttersprache im Sprachorientierungsunterricht; vgl. Frank, 1976.B.1) oder u.U. (vor allem in der Erwachsenenbildung) auch durch verbales Beschreiben dieser Struktur, womit Lm zu einem sprachlichen Modell entartet.

3. *Entwicklung eines kybernetisch-pädagogischen Transferbegriffs*

Von (positivem) „Transfer" (Übertragung) eines Lernstoffes $L1$ (z.B. die Internacia Lingvo) auf einen späteren Lernstoff $L2$ (z.B. eine fremde Nationalsprache wie Englisch, Französisch usf.) spricht man in der Lernpsychologie, wenn eine Lernergruppe, die zuvor $L1$ erlernt hatte, in derselben Zeit einen höheren Lernzustand (oder denselben Lernzustand in kürzerer Zeit) erreicht als eine sonst gleichartige Lernergruppe ohne vorherige Aneignung von $L1$ (Arnold, Eysenck, Meili, 1972). Für den breiten Anwendungsbereich der Lernsteuerung bzw. der Klassenschulung (die Voraussetzungen sind in Frank, 1977.A.1 diskutiert) wurde die Lernerfolgsfunktion

$$(10) \qquad p(t) = 1 - (1-p_o) \cdot e^{-\eta C_v t/I}$$

aufgestellt, aus welcher die Lernzeitformel (1) durch Auflösung nach t hervorgeht. Da C_v ein altersabhängiges Adressatenmerkmal ist, kann eine Vergrößerung von $p(t)$ in konstanter Lernzeit (bzw. eine Verkleinerung von t bei konstantem Lehrziel $!p$) offenbar nur erreicht werden, wenn man wenigstens eine der drei folgenden Veränderungen bewirkt:

- a) man vergrößert η
- b) man vergrößert p_o
- c) man verkleinert I.

Die *Effikanz* wurde als ein Merkmal des Lehrsystems und der Unterrichtsumwelt definiert, also als adressatenunabhängige Größe, die daher nicht transferbegründend wirkt. Durch eine spätere Verallgemeinerung (Frank, 1977.A.2) wurde jedoch versucht, die Motivlage der Lerner zu berücksichtigen. Wird diese durch vorherige Aneignung von $L1$ verbessert (was in unserem sprachpädagogischen Beispiel als unbestrittene, aber bisher nicht quantifizierte Erfahrungstatsache gilt), dann wird Transfer bewirkt. Allerdings ist auch die „subjektive Effikanz" nach oben grundsätzlich durch 1 beschränkt!

Die *Vorkenntnis* stellt die nächstliegende Deutung des Transfers dar, kann jedoch allein die Beziehung (3) nicht begründen. Ist nämlich in $L1$ schon der Prozentsatz p_o der Lehrstoffinformation $I(L2)$ enthalten, dann müßte für $L1$ mindestens die für das Erlernen von $p_o \cdot I(L2)$ nötige Zeit aufgewendet werden, d.h. die Kurve $p(o;\frac{\eta}{I})$ des Lernerfolgs wird nur entsprechend vorverlegt (und damit zur Kurve $p(p_o;\frac{\eta}{I})$ unseres Bildes.

Die Reduktion der *Lehrstoffinformation* $I(L2)$ auf den Bruchteil $r<1$ nach der Beziehung

$$r = \frac{I(L2|L1)}{I(L2)}$$

führt (ebenso wie die Vergrößerung von η, jedoch ohne vergleichbare Beschränkung!) zu einem steileren Anfangswachstum der Lernerfolgsfunktion (hypothetische Kurve $p(o;\frac{\eta}{rI})$ und erwartete Kurve $p(p_o;\frac{\eta}{rI})$ im Bild).

Da nicht auszuschließen ist, daß auch ohne vorherige Aneignung von $L1$ bereits eine Vorkenntnis p_{oo} besteht, ergibt sich für den durch Zeitgewinn gemessenen Transfer aus (1)

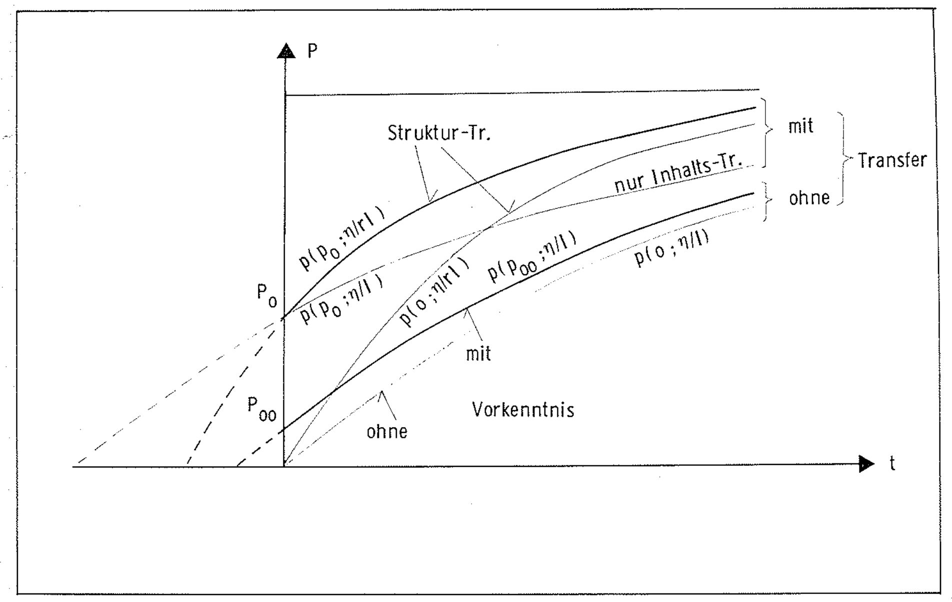
$$\Delta t = t_o - t = \frac{I}{\eta C_v}\left(\ln\frac{1-p_{oo}}{1-!p} - r\cdot\ln\frac{1-p_o}{1-!p}\right) \qquad (11)$$

$$= \frac{I}{\eta C_v}\ \ln\frac{1-p_{oo}}{(1-p_o)^r\cdot(1-!p)^{1-r}}$$

wobei t_o die Zeit bezeichnet, welche ohne Transfer zur Erreichung des Lehrziels $!p$ benötigt wird, t die (kürzere) Zeit aufgrund des (positiven) Transfers. Der prozentuale Zeitgewinn ist

$$\frac{\Delta t}{t_o} = 1 - r\cdot\frac{\ln(1-p_o)-\ln(1-!p)}{\ln(1-p_{oo})-\ln(1-!p)} \qquad (12)$$

Auch dieses Transfermaß hängt nicht nur von r und p_o ab, sondern darüber hinaus vom erstrebten Lehrziel $!p$; dieses könnte (willkürlich) standardisiert werden, z.B. durch



Auswirkung des inhaltlichen bzw. des strukturellen Transfers auf die Lernerfolgskurve

$!p = 1-1/e \approx 63\,\%$. Dann wird

$$\left.\frac{\Delta t}{t_o}\right|_{!p\,=\,1-1/e} = 1-r\cdot\frac{1+\ln(1-p_o)}{1+\ln(1-p_{oo})} = 1-r\cdot\frac{1-p_o-\sum_{n=2}^{\infty}\frac{p_o^n}{n}}{1-p_{oo}-\sum_{n=2}^{\infty}\frac{p_{oo}^n}{n}} \qquad (12a)$$

wobei für Vorkenntniswerte p_o bzw. p_{oo} unter 20 % die unendlichen Summen vernachlässigbar werden.

Eine andere, vielleicht natürlicher erscheinende Möglichkeit, die Transferwirkung durch eine skalare Maßgröße zu erfassen, bezieht sich auf den Fall des Unterrichtens von $L2$ durch eine ideale Lernregelung (vgl. Frank, 1977.A.5), also auf die theoretisch wirksamste Unterrichtsform. Man kann dabei den transferbedingten prozentualen Zeitgewinn bis zur *vollständigen* Aneignung von $L2$ bei einer solchen (individuellen) Unterrichtsform als Transfermaß einführen. Im Falle der idealen Lernregelung (ILR) ist die erforderliche Lernzeit ohne bzw. mit Transferwirkung offenbar

(13a) $t_o^{ILR} = \frac{(1-p_{oo}) \cdot I}{\eta \cdot C_v}$ bzw. (13b) $t^{ILR} = \frac{(1-p_o) \cdot r \cdot I}{\eta \cdot C_v} \leqslant t_o^{ILR}$

und damit die prozentuale Hilfe h, die $L1$ für die vollständige Erlernung von $L2$ hinsichtlich der erforderlichen Lernzeit darstellt:

$$(14) \qquad h =_{Df} \frac{t_o^{ILR} - t^{ILR}}{t_o^{ILR}} = 1 - r \frac{1-p_o}{1-p_{oo}}$$

Dieses Ergebnis ist ein Argument für die Standardisierung, die zu (12a) führte. Da p_o, p_{oo} zwischen 0 und 1 liegen, wird h keinesfalls größer als 1. Jedoch ist andererseits das Maß geeignet, auch negativen Transfer zu erfassen, denn für $p_o < p_{oo}$ oder $r > 1$, also im Falle inhaltlicher oder struktureller *Lernerschwerung,* kann h negativ werden. Diese unerwünschten Eigenschaften des eingeführten Maßes h dürfen nicht den Umstand verschleiern, daß man von den drei erwähnten Transferursachen zwar die *motivationale* und die *strukturelle* durch eine *gemeinsame Zahl r* erfassen kann, solange man nur Zusammenhänge betrachtet, in welche I und η nur als Quotient η / I eingehen, daß aber die transferbedingte *inhaltliche* Vorkenntnis p_o für die Anwendung von (11) *zusätzlich* bekannt sein muß. Von den drei Größen p_o, r, h werden also in unserem Zusammenhang zwei benötigt, d.h. eine theoretisch befriedigende Beschreibung des Transfer nur durch eine skalare Größe gelingt nicht.

4. *Anwendungsmöglichkeit der informationstheoretischen Transferanalyse*

Die Bestimmung von p_o und p_{oo} zu Beginn des Unterrichts über $L2$ kann in traditioneller Weise durch einen Stichprobentest erfolgen. Zur Messung von r vereinigt man die beiden Lernergruppen zu einer gemeinsam über $L2$ zu unterrichtenden Klasse, damit sichergestellt ist, daß η für beide übereinstimmt (bzw. daß ein Unterschied in diesem Parameter adressatenseitig bedingt ist, also die motivationale Transferkomponente betrifft, die wir gemeinsam mit der strukturellen durch r erfassen können). Nach der (theoretisch willkürlich wählbaren) Unterrichtsdauer T wird der Lernzustand erneut gemessen. Er betrage p_T bei den Lernern, welche durch den Transfer von $L1$ begünstigt sind, p_{oT} bei den ohne Transferbegünstigung lernenden Adressaten des gemeinsamen Unterrichts. Setzt man in beiden Fällen den erreichten Zustand statt $!p$ in (1) ein, dann erhält man wegen der übereinstimmenden Zeit $t = T$

$$(15) \qquad \frac{I}{\eta C_v} \ln \frac{1-p_{oo}}{1-p_{oT}} = \frac{r \cdot I}{\eta C_v} \ln \frac{1-p_o}{1-p_T}$$

und daraus

$$(16) \qquad r = \frac{\ln(1-p_{oo}) - \ln(1-p_{oT})}{\ln(1-p_o) - \ln(1-p_T)} = \frac{\ln w_o}{\ln w}$$

Da r nur vom Lernerfolgsmaß

$$(17) \qquad w = \frac{1-p_o}{1-p_T}$$

abhängt, kann die Messung unbeschadet der Wahrscheinlichkeit zufällig richtiger Ankreuzungen sogar mit demselben Auswahlantworttest im Vor- und Nachtest durchgeführt werden (Frank, 1977.A.5). Man kann auch in anderer, vor und nach Unterricht übereinstimmender Weise die Lehrstoffbeherrschung stichprobenhaft prüfen und nur die jeweiligen Fehler F_o, F_T, F_{oo}, F_{oT} zählen. Hätten maximal N Fehler gemacht werden können (eine Zahl, die oft – z.B. bei einem Diktat oder einer Übersetzung – schwer definierbar ist!), dann ist offenbar allgemein

$$(18) \qquad p = \frac{N-F}{N} = 1 - \frac{F}{N}$$

so daß aus (17) folgt

$$(19) \qquad w = \frac{F_o}{F_T}$$

und aus (16)

$$(20) \qquad r = \frac{\ln F_{oo}/F_{oT}}{\ln F_o/F_T}.$$

5. *Sprachpädagogische Anwendung*

Trotz der Bemühungen von Landa (1966 u.a.), Bung (1972) und Zierer (1970) ist ein bisher befriedigender Aufbau einer prospektiven (d.h. auf Vorhersagen abzielenden) Fremdsprachpädagogik auf der Basis der kybernetischen Pädagogik allenfalls bruchstückhaft (vgl. Meder, 1978) gelungen. Offensichtlich bereitet die Vielschichtigkeit und Verwickeltheit fremdsprachlicher Lehrstoffe einer kybernetisch-pädagogischen Aufarbeitung derzeit noch große Schwierigkeiten. Für die Anwendung der cartesischen Methode fehlte bisher ein hinreichend einfacher Fall dieses Problemkreises, mit dem begonnen und dabei die Forschungs- und Entwicklungsmethode weiterentwickelt werden könnte. Einen solchen Einstieg bietet nun aber die von der deutschen Philologie bisher vernachlässigte Interlinguistik, die einen Brückenschlag von der kybernetischen Pädagogik zur Sprachpädagogik liefert. Auf dem in der gegenwärtigen Arbeit skizzierten theoretischen Fundament wird nämlich die als Lehrstoffmodell vorgeschlagene Internacia Lingvo (ILo) hinsichtlich ihrer Transferqualität aufgrund empirischer Messungen beurteilbar. Eine sprachwissenschaftliche Analyse ergibt, daß die ILo auf den verschiedenen Superierungsstufen der Sprache als einfaches, regelmäßiges und deutliches Fremdsprachmodell, folglich als geeignetes Lehrstoffmodell anzusehen ist:

(1) Auf der Stufe der *Phoneme* enthält die ILo als Modell von gesprochenen *und* geschriebenen (also von „Kultur"-)Sprachen einen Vorrat von Schriftzeichen (*Buchstaben*). *Unabhängig* von seinen jeweiligen Vorgängern oder Nachfolgern im Gesamttext (Deutlichkeit!) ist jedes der zu diesem Alphabet gehörigen *nur* 28 lateinischen Lautschriftzeichen (Einfachheit!) *eindeutig* (Regelmäßigkeit!) seiner Aussprache zugeordnet.

(2) Die kleinsten Laut- bzw. Schriftzeichenfolgen, die für eine Sinneinheit (ein *Semantem*) stehen, heißen *Moneme* oder Wortteile, das sind Lexeme (Wortwurzeln und Affixe) oder Morpheme (grammatische Endungen). Die Zahl der nicht fachspezifischen Moneme liegt bei nur etwa 2000 (Einfachheit!), die kontextunabhängig (Deutlichkeit!) und ausnahmefrei (Regelmäßigkeit!) stets demselben Semantem zugeordnet sind.

(3) Der Ausdrucksreichtum der ILo wird durch Kombination der Moneme zu einem offenen Vorrat von *Wörtern* erzeugt, die nicht gelernt werden brauchen, da ihre Synthese bzw. Analyse algorithmisch erfolgen kann. (Nur in wenigen Fällen führt die Analyse zu Mehrdeutigkeiten, die kontextabhängig aufzulösen bzw. für Wortspiele auszunutzen sind.)

(4) Wie in manchen Nationalsprachen ist auch in der ILo ein zusammengesetzes Wort zu einer Wortgruppe (*Syntagma*) äquivalent (Hundezüchter = Züchter von Hunden; instrucelkonforma = konforma al la instrucelo = konforma al la celo de la instruo).

Die auf dieser Stufe und auf der anschließenden Stufe

(5) des *Satzes* (der Kombination von Syntagmen) geltenden, also zu erlernenden, Restriktionen der möglichen Kombinationen aus Unterzeichen sind bei der Ilo der Zahl nach sehr geringfügig („lingvo internacia" = „internacia lingvo", „knabo vidis rozeton" = „vidis knabo rozeton" = „rozeton vidis knabo", aber „nur paroli internacie" ≠ „paroli nur internacie").

Eine kybernetische Interlinguistik sollte versuchen, die in (5a) auftauchenden Größen für y = ILo und z = Englisch, Französisch, Latein usf. durch Feinanalysen auf den verschiedenen Stufen abzuschätzen. Inwieweit (d.h. genauer: für welche Zielausprägungen $!p$) die entscheidende Ungleichung (3) erfüllt wird, läßt sich ermitteln, wenn man die Dauer $Z(Lm)$ des Sprachorientierungsunterrichts, die Werte I, η, C_v und p_{oo} des späteren Unterrichts einer Nationalfremdsprache sowie die Transferwerte r, p_o kennt: die aus (11) als Funktion von $!p$ berechenbare Zeitersparnis muß größer als $Z(Lm)$ sein. Die Lehrstoffinformation $I(Lg)$ kann aus der leichter ermittelbaren Lehrstoffinformation des zu Lg äquivalenten Teils der ILo durch Multiplikation mit der relativen Schwierigkeit bestimmt werden (Frank, 1976.A.3., 1977.A.3).

Nur um eine erste, grobe Vorstellung von den Größenordnungen der in der Transinformationsanalyse auftretenden Zahlen zu erhalten, verweisen wir auf Szerdahelyi (1970), nach welchem ein Englischunterricht für ungarische Kinder zu 60% erfolgreich war

und die Lernerleichterung für diese Kinder durch vorhergehendes Erlernen der ILo 40% betrug. Diesen letzteren Wert setzen wir für $\Delta t / t_o$ in (12) ein. Da 60% ≈ 63%, benutzen wir die Beziehung (12a) und erhalten näherungsweise

$$0{,}4 = 1 - r \cdot \frac{1-p_o}{1-p_{oo}} .$$

Schätzt man die von Szerdahelyi nicht gemessenen (aber sicher unter 20% liegenden) Werte p_o und p_{oo} auf 10% bzw. 5% (das Ergebnis r hängt davon nur unerheblich ab!), dann kommt man auf einen Strukturtransfer der ILo auf die englische Sprache von $r \approx 0{,}63$.

Eine solche Vorausabschätzung des Ergebnisses noch laufender, umfangreicher empirischer Untersuchungen kann natürlich nur den Sinn einer Orientierungshilfe z.B. bei der Festlegung eines meßtechnisch geeigneten Zeitpunkts T für die Messung p_T und p_{oT} (nämlich zu einer Zeit theoretisch großer Differenz dieser Werte!) haben. —

Schrifttum

Arnold, Eysenck, Meili (Hrsg.): Lexikon der Psychologie, Herder, Freiburg/i.Br. 1972

Bung, Klaus (1972): Ansätze zu einer Theorie des Programmierten Sprachunterrichts. Julius Groos Verlag, Heidelberg, 1972

Frank, H. (1959.N.1 und 1968.N.1): Grundlagenprobleme der Informationsästhetik und erste Anwendung auf die mime pure (Dissertation, TH Stuttgart), 1. Auflage: Hess, Waiblingen, 1959, 2. Auflage: Schnelle, Quickborn, 1968, Nachdruck in Meder (1973/74.W.1), Bd. 5

Frank, H. (1960.A.3): Über das Intelligenzproblem in der Informationspsychologie. In: GrKG 1960, Bd. 1, H. 3, S. 85—96 (Nachdruck in Meder, 1973/74.W.1, Bd. 1)

Frank, H. (1961.A.2): Zur Mathematisierbarkeit des Ordnungsbegriffs. In: GrKG 1961, Bd. 2, H.2, S. 33—42. (Nachdruck in Meder, 1973/74.W.1, Bd. 1)

Frank, H. (1969.0.1): Kybernetische Grundlagen der Pädagogik. Eine Einführung in die Pädagogistik für Analytiker, Planer und Techniker des didaktischen Informationsumsatzes in der Industriegesellschaft. 1. Band: Allgemeine Kybernetik (XX+409 S.), 2. Band: Angewandte kybernetische Pädagogik und Ideologie (X+290 S.). Baden-Baden: Agis. Stuttgart: Kohlhammer, 2. Aufl. 1969

Frank, H. (1974.A.3): Ein Ansatz zu einer kybernetisch-pädagogischen Lehrplanungstheorie. In: Neue Unterrichtspraxis, 1974, H.6, S. 340—347 und (mit Berichtigungen) in G. Lobin (Hrsg.), Kybernetik und Bildung III, Schroedel, Hannover, 1976

Frank, H. (1975.A.5): Lehrwirkungsgrad und Lernzeit. In: GrKG 1975, Bd. 16, H.4, S. 113—120

Frank, H. (1975.B.1): Ein Ansatz zu einer kybernetisch-pädagogischen Lehrplanungstheorie als Beitrag zur Curriculumdiskussion. In: K. Frey u.a. (Hrsg.): Curriculum-Handbuch, München: R. Piper & Co., 1975, S. 486—492

Frank, H. (1976.A.1): Über den informationspsychologischen Zusammenhang zwischen dem Wirkungsgrad eines Unterrichts und dessen Lehrziel. In: G. Lobin (Hrsg.): Kybernetik und Bildung II. Paderborner Werkstattgespräche Bd. 8. Hannover/Paderborn: Schroedel/ Schöningh, 1976, S. 151—161

Frank, H. (1976.A.3): Zur relativen Lernleichtigkeit einiger Sprachen. In: GrKG 1976, Bd. 17, H.4, S. 120—124

Frank, H. (1976.B.1): Sprachorientierungsunterricht nach dem Paderborner Modell. In: aula 1976, H.2, S. 133—141

Frank, H. (1976.H.1): Mallonga Enkonduko en la Kibernetikan Pedagogion. In: H. Behrmann, S. Stimeč (Hrsg.): Klerigo kaj Prikalkulado. Europäische Reihe: Entnationalisierte Wissenschaft. Bd. 1, Bamberg: difo-druck 1976, S. 9—50

Frank, H. (1972.H.2): La lerntempo enplanigenda depende de la instrusistemo. In: S. Jarmark und F. Januszkiewicz (Red.): Technologia Kształcenia. Zakład Graficzny Politechnik Poznanskiej. Posen, 1976, S. 586–591
Frank, H. (1977.A.1): Die Lehrerfolgs- und Zeitbedarfsprognose mit dem β-η-Diagramm. In: GrKG 1977, 18, H.2, S. 45–56
Frank, H. (1977.A.2): Derzeitige Bemühungen um Erweiterungen des informationspsychologischen Modells. In: GrKG 1977, Bd. 18, H.3, S. 61–71
Frank, H. (1977.A.3): Die Meßmöglichkeit der relativen Schwierigkeit von Fremdsprachen. In: G. Lobin, W.D.E. Bink (Hrsg.): Kybernetik und Bildung III. Paderborner Werkstattgespräche Bd. 11. Hannover/Paderborn: Schroedel/Schöningh 1977, S. 14–25
Frank, H. (1977.A.5): Begriff, Eigenschaften und Anwendung des Bildungsinkrements als Maß des Lernerfolgs. In: GrKG 1977, 18.4, S. 103–112
Frank, H. (1978): Zur Behebung des Widerspruchs zwischen traditioneller Fremdsprach-Lehrplanung und europäischer Kultur- und Gesellschaftspolitik. In: A. Melezinek (Hrsg.), Technische Medien im Sprachunterricht, Leuchtturm-Verlag, Konstanz, 1978, S. 57–60
Hilgers, R. (1978): Zur Deduktion der Lernzeitformel aus dem diskreten ALZUDI-Modell. GrKG 19/2, 1978, S. 33–44
Landa, L.N. (1966): Diagnostik und Programmierter Unterricht. In: H. Frank (Hrsg.): Lehrmaschinen in kybernetischer und pädagogischer Sicht, Bd. 4, Stuttgart-München: Klett-Oldenbourg, 1966, Bd. 4, S. 54–67
Lánský, M. (1971): Weiterentwicklung der Methode VERBAL – Methodischer Hinweis für die Vorbereitung des Lehrstoffes. In: Rollett und Weltner (Hrsg.): Fortschritte und Ergebnisse der Unterrichtstechnologie, München, Ehrenwirth 1971, S. 118–125
Lánský, M. und Polak, V. (Hrsg.): Studien zur Superierung durch Komplexbildung. Schöningh, Paderborn und Schroedel, Hannover 1977
Meder, B.S. (1978.C.1): Prolegomena zu einer kybernetisch orientierten Fremdsprachpädagogik. Paderborner Arbeitspapier Nr. 33. Institut für Kybernetische Pädagogik. Paderborn: FEoLL GmbH, 1978, 67 S.
Meder, B.S. (1973/74.W.1) und Schmid, W. F. (Hrsg.): Kybernetische Pädagogik, Schriften 1958–1972. Dünndruck-Quellensammlung. Kohlhammer, Stuttgart-Berlin-Köln-Mainz, Bd. 1–4, 1973, Bd. 5, 1974
Stachowiak, H. (1973.N.1): Allgemeine Modelltheorie. Wien-New York: Springer 1973
Szerdahelyi, I. (1970: La didaktika loko de la Internacia Lingvo en la sistemo de lernejaj studobjektoj. In: Internacia Pedagogia Revuo, Heft 0, 1970, Verlag Ludwig Pickel, Libroservo
von Cube, F. (1961.A.1): Über ein Verfahren der mechanischen Didaktik. In: GrKG 1961, Bd. 2, H.1, S. 7–10 (Nachgedruckt in Lánský u. Polák, 1977)
von Cube, F. (1970.N.1): Kybernetische Grundlagen des Lernens und Lehrens. Stuttgart: Klett, 3. Aufl. 1970
Weltner, K. (1970.N.1): Informationstheorie und Erziehungswissenschaft. Quickborn: Schnelle, 1970
Weltner, K. (1975.A.1): Über ein Verfahren zur Bestimmung der optimalen Reihenfolge für Lehrstoffanordnung. In: G. Lobin (Hrsg.): Kybernetik und Bildung I. Schöningh/Schroedel, Paderborn/Hannover, 1975, S. 104–113
Weltner, K. (1977): Kognitive Verlaufsfunktion – Zum Abbau der subjektiven Information von regelhaften Zahlenfolgen und geometrischen Mustern. Grundlagenstudien aus Kybernetik und Geisteswissenschaft Bd. 18, H 3, 1977, S. 79–83
Zierer, Ernesto (1970): Elementos de Pedagogia Cibernética para la Didáctica de los Idiomas Extranjeros. Universidad Nacional de Trujillo, Trujillo, 1970

Eingegegangen am 16. August 1978

Anschrift des Verfassers:
Prof. Dr. Helmar Frank, Rolandsgärten 2, D-4790 Paderborn

Notiz zur Zufallskorrektur bei Tests mit Ratetreffern

von Rainer HILGERS, Paderborn
aus dem FEoLL-Institut für Kybernetische Pädagogik, Paderborn (Direktor: Prof. Dr. Helmar Frank)

1. *Problemstellung und klassischer Lösungsansatz*

Bei Tests, in denen Antworten nur in gebundener Form abzugeben sind, kann die Zahl R der richtigen Lösungen größer sein als nach dem tatsächlichen Wissen eines Probanden zu erwarten wäre. Werden n Fragen gestellt, so führt die (z.B. durch einen voraufgegangenen Unterricht erworbene) Kompetenz p des Adressaten im Mittel zu pn richtigen Antworten; darüber hinaus können weitere Lösungen erraten werden. Für viele Zwecke ist es aber erwünscht, über eine Formel zu verfügen, die aus dem beobachteten (Zufalls-)Wert R auf p schließen läßt. Um den mathematischen Aufwand auf einem vertretbaren Niveau zu halten, wird man annehmen müssen, daß der Adressat bei allen Kenntnislücken mit gleicher Wahrscheinlichkeit z zufällig richtig rät. Handelt es sich z.B. um einen Test mit Auswahlantworten, so unterstellt man
(a) eine Menge von informationell unabhängigen Testfragen
(b) zu jeder Frage die gleiche Zahl m von Distraktoren (Antwortalternativen)
(c) deren subjektive Gleichwahrscheinlichkeit für einen Unkundigen.

Unter diesen Voraussetzungen darf man

$$(1) \qquad z = \frac{1}{m}$$

setzen.

Wenn sicher ist, daß der Proband in allen Fällen von Unkenntnis einen Rateversuch unternommen hat, dann enthält sein Testergebnis im Mittel $z(1-p)n$ „Ratetreffer" und es gilt

$$(2) \qquad E(R) = n \cdot [p + (1-p)z] .$$

Bezeichnet man mit $r =_{Df} R/n$ den A n t e i l richtig gelöster Aufgaben, so folgt also

$$(3) \qquad \bar{r} = E(r) = p + (1-p)z$$

und damit

$$(4) \qquad p = \frac{\bar{r}-z}{1-z} .$$

Es ist allgemein üblich (so z.B. bei Frank, 1977), die Formel (4) zur Schätzung der Kompetenz des Adressaten zu verwenden, indem man auf der rechten Seite $\bar{r}$ durch r ersetzt:

$$\hat{p} = \frac{r-z}{1-z}. \tag{5}$$

Lienert (1969) nennt $X =_{\mathrm{Df}} pn$ den Rohwert des Probanden und empfiehlt ohne weitere Begründung bei F Fehlern in einem Test mit Mehrfachwahlaufgaben die Korrektur

$$X = R - \frac{F}{m-1} \tag{6}$$

(a.a.O., S. 83). Diese Formel ist wegen (1) mit (5) äquivalent. Genaugenommen bedeutet der hier vorgeschlagene Weg, daß aus dem Ergebnis eines e i n z i g e n Experiments auf den Mittelwert einer unbekannten Verteilung geschlossen wird — ein logisch wie statistisch unbefriedigendes Verfahren. Die in Formel (5) verborgene Fehlerquelle wird besonders deutlich im Sonderfall $r = 1$, woraus $p = 1$ zu folgern wäre. Die numerische Aussage dieser Konklusion kann zwar hin und wieder zutreffend sein, stellt aber auf keinen Fall das statistische Mittel dar (wenn man einem Lottogewinner nicht die vorherige Kenntnis der Gewinnzahlen unterstellen will).

Für die Problematik, daß statistische Gesetze ihren Sinn verlieren können, wenn man abhängige und unabhängige Variablen vertauscht, findet man andernorts (Hilgers, 1978) schon ein vom Autor durchgerechnetes bildungswissenschaftliches Beispiel. Seinerzeit konnten jedoch die für den Rückschluß maßgeblichen Wahrscheinlichkeiten ebenfalls aus dem zugrundeliegenden Modell deduziert werden. Der hier diskutierte, qualitativ andersartige Fall scheitert nicht an der Frage der Zulässigkeit einer Transformation von (3) in (4), sondern daran, daß die Verteilungsfunktion $P(p \leqslant x)$ nicht durch den meßbaren Erfolgsanteil r determiniert ist. Die gestellte Aufgabe entstammt also dem Bereich der beurteilenden Statistik, wo die Schätzung eines verborgenen Parameters aus einer begrenzten Zahl empirischer Daten in vertretbarer Form zu geschehen hat. Die vorliegende Arbeit setzt sich zum Ziel, ausgehend von zwei sehr verschiedenen Ansätzen zu zeigen, daß die Schätzung nach Formel (5) zwar nicht im strengen Sinne „richtig", aber unter den vorausgesetzten Umständen die bestmögliche ist.

2. *Eine Näherungsrechnung vermittelt durch ein Urnenmodell*

Die Darstellung in diesem Abschnitt folgt fast ausnahmslos einem bisher unveröffentlichten Manuskript, welches wir Hengst (1978) verdanken. Hengst bemerkt, daß die Zahl der mißlungenen Rateversuche eine e m p i r i s c h e Größe $(n-R)$ ist. Er zeigt, wie man daraus auf die wahrscheinliche Gesamtzahl U von Rateversuchen schließen kann, wenn unbegrenzt viele Aufgaben gestellt sind und wenn d e r l e t z t e F e h l e r z u g l e i c h d e r l e t z t e R a t e v e r s u c h w a r. (Der approximative Charakter dieser Betrachtungsweise beruht also hauptsächlich auf der Vernachlässigung der Rolle von $n!$) Die zweite der genannten Bedingungen hat zur Folge, daß unsere Argumentation den Fall $R = n$ nicht einschließt.

Man stelle sich eine Urne vor, die mit roten und schwarzen Kugeln im Verhältnis $z:(1-z)$ gefüllt ist. Zieht man mit Zurücklegen, dann kann man jede Ziehung als Rateversuch, jede schwarze Kugel als falsch geratene Antwort interpretieren. Wie lange wird man durchschnittlich ziehen müssen, bis man die erste schwarze Kugel in der Hand hält? Die Wahrscheinlichkeit, mit genau k roten Kugeln zu beginnen, ist gleich $z^k(1-z)$. Die Wartezeit auf eine schwarze Kugel beträgt also im Mittel

$$\begin{aligned} E(T_1) &= \sum_{\mathbb{N}} (k+1) z^k (1-z) = (1-z) \sum_{\mathbb{N}} \frac{d}{dz} z^{k+1} \\ &= (1-z) \frac{d}{dz} \sum_{\mathbb{N}} z^{k+1} \\ &= (1-z) \frac{d}{dz} \left(\frac{z}{1-z}\right) = \frac{1}{1-z} \end{aligned} \tag{7}$$

Da es sich um einen Markoffprozeß handelt, ist die durchschnittliche Dauer des Experiments proportional zu der gewünschten Zahl schwarzer Kugeln, also

$$E(U) = E(T_{n-R}) = \frac{n-R}{1-z}. \tag{8}$$

Falls n genügend groß ist, scheint $E(U)$ als Näherung für den u n b e k a n n t e n Mittelwert μ der Verteilung von U in einer e n d l i c h e n Zahl von Aufgaben geeignet. Damit wäre (vgl. (5))

$$\begin{aligned} E(p) &= E(pn)/n = \frac{1}{n}[n-\mu] \\ &\approx \frac{1}{n}[n-E(U)] \\ &= \frac{1}{n}\,\frac{R-nz}{1-z} = \frac{r-z}{1-z}. \end{aligned} \tag{9}$$

3. *Ableitung einer Maximum-Likelihood-Schätzfunktion für p*

Nach der M e t h o d e d e r m a x i m a l e n M u t m a ß l i c h k e i t (Maximum-Likelihood-Prinzip) wählt man als Schätzung eines unbekannten statistischen Parameters denjenigen Wert, der das empirische Meßergebnis am wahrscheinlichsten macht. Probanden mit der Kompetenz p antworten mit der Wahrscheinlichkeit

(10) $h = z + (1-z)p$

richtig auf eine Testfrage. Statt nach einer Schätzfunktion für p kann man auch nach einer Schätzfunktion für h suchen. Die Wahrscheinlichkeit eines bestimmten Testprotokolls mit R richtigen Lösungen beträgt

(11) $l = h^R (1-h)^{n-R}$

Rechentechnisch ist es vorteilhafter, statt l die Funktion $\ln l = R \ln h + (n-R) \ln(1-h)$ zu betrachten und deren Maximum in Abhängigkeit von p zu bestimmen. Es gilt

$$(12) \quad \frac{\partial \ln l}{\partial p} = \frac{R}{h}\frac{\partial h}{\partial p} + (n-R)\frac{1}{1-h}(-1)\frac{\partial h}{\partial p}$$

$$= \frac{\partial h}{\partial p}\left[\frac{R}{h} - \frac{n-R}{1-h}\right]$$

$$= (1-z)\frac{R-nh}{(1-h)h}$$

$$= \frac{R-nh}{(1-p)h}$$

Die rechte Seite verschwindet für $R-nh = 0$. Demnach wäre

$$(13) \quad \hat{h} = \frac{R}{n} = r$$

die Maximum-Likelihood-Schätzfunktion für h, womit wegen (10) wieder (5) herauskommt.

Schrifttum

Frank, H.: Begriff, Eigenschaften und Anwendung des Bildungsinkrements als Maß des Lernerfolgs, in: GrKG 18/4, 1977, S. 105–112
Hengst, M.: Statistische Aspekte der Auswertung und Beurteilung von Prüfungen und Leistungstests. Unveröffentlichtes Manuskript, Berlin 1978
Hilgers, R.: Zur Deduktion der Lernzeitformel aus dem diskreten Alzudi-Modell, in: GrKG 19/2, 1978, S. 33–43
Lienert, G. A.: Testaufbau und Testanalyse. Beltz, Weinheim, 3. Auflage 1969

Eingegangen am 29. Juli 1978

Anschrift des Verfassers:
Dr. Rainer Hilgers, Erwin-Rommel-Straße 24, D-4790 Paderborn

Eine europäische Fallstudie zum Zusammenhang zwischen Information und Entscheidungsfreiheit

von Brigitte S. MEDER, Paderborn
aus dem FEoLL-Institut für Kybernetische Pädagogik Paderborn (Direktor: Prof. Dr. Helmar Frank)

1. *Problemstellung*

Zwei Gruppen von Schülern, von denen die eine zusätzlich zum obligatorischen Englischunterricht an einem zweijährigen Plansprachunterricht teilnahm (wir bezeichnen diese Gruppe mit PL) sollten hinsichtlich ihres Interesses an Ländern der Europäischen Gemeinschaften (EG) miteinander verglichen werden. Es ist anzunehmen, daß Interessen hervorgerufen werden durch bestimmte Merkmale der Gegenstände dieser Interessen, z.B. ist das Interesse an einem bestimmten Land bedingt durch dessen Bewohner, durch die dort gesprochene Sprache, die dortigen Speisen u.s.w. Die Unterschiede dieser Merkmale sind in diesem Beispiel die Ursache für die Bevorzugung eines Landes oder einiger Länder gegenüber anderen.

Man kann mit D. und A. Gabor (1954) der Meinung sein, daß sich die Freiheit der Wahl in der Verschiedenheit der effektiv getroffenen Entscheidungen bei einer größeren Anzahl von sich Entscheidenden ausdrückt. Aber man muß darüber hinaus auch davon ausgehen, daß eine maximale Entscheidungsfreiheit nur bei denen gegeben ist, die eine genaue Kenntnis über alle Entscheidungsmöglichkeiten haben, m.a.W. über die Merkmale aller zur Wahl stehenden Entscheidungsmöglichkeiten gleichmäßig – ideal – informiert sind. Wer nämlich die Vorzüge nur eines Teils der wählbaren Gegenstände eines Interesses kennt, über Existenz und Qualität der anderen Möglichkeiten aber uninformiert ist, der wird diese kaum einmal wählen: er entscheidet nicht (durch Information) *emanzipiert*, sondern (durch Informationsvorenthaltung) *manipuliert*. Die für die zu treffenden Entscheidungen in unserer Untersuchung (also für die Wahl zwischen einzelnen Ländern) relevanten Kenntnisse werden teilweise schon im Englischunterricht aber in noch größerem Maße im Plansprachunterricht vermittelt, so daß zu erwarten ist, daß der Informiertheitsgrad der Plansprachschüler (PL) näher dem des „Ideal-Informierten" ist als der der „Nur-Englisch-Schüler" (N-PL); m.a.W. es ist zu erwarten, daß die Entscheidungsfreiheit der Plansprachschüler größer ist als die der Nicht-Plansprachschüler. Das hätte zur Folge, daß bei der N-PL-Gruppe ein deutlich merkbarer einseitiger Anstieg des Interesses für einige wenige Länder beobachtet werden müßte gegenüber einer gleichmäßigeren Verteilung der Interessensbekundungen der PL-Gruppe auf alle zur Wahl angebotenen Entscheidungsmöglichkeiten.

Die vorliegende Untersuchung beschreibt und analysiert so geartete empirische Beobachtungen.

2. *Untersuchungsmethode*

Wir vergleichen bei zwei Gruppen von Versuchspersonen (Vpn) *nicht* das *Ausmaß* des Interesses an einem bestimmten Land, sondern zunächst lediglich die *Entscheidung* für ein Land oder mehrere Länder aufgrund irgendwie gearteter und durch eine Frage angesprochener Merkmale. Dazu benötigten wir nicht eine z.B. nach dem Thurstone-Verfahren (Thurstone, 1945; Hofstätter, 1966, S. 165f) konstruierte abgestufte Skala, sondern zunächst lediglich einige in Fragen angesprochene Merkmale, die mit großer Wahrscheinlichkeit beim Entstehen des Interesses an Ländern beteiligt sind.

Wir formulierten folgende 14 Fragen:

1. In welches Land würdest Du gerne eine Klassenfahrt machen?
2. In welchem Land würdest Du gerne Deine Ferien verbringen?
3. Über welches Land würdest Du Dir gerne ein Buch ausleihen?
4. Die Sagen und Märchen aus welchem Land interessieren Dich?
5. Welche Witze interessieren Dich, die Witze über die Belgier, die Dänen u.s.w. ...
6. Wen würdest Du gerne als Austauschschüler zu Deinen Eltern nach Hause einladen?
7. Über welches Land würdest Du Dir gerne einen Film ansehen?
8. Im Radio oder im Fernsehen wird ab und zu etwas über fremde Länder berichtet. Wenn Du Dir das Land, über das berichtet wird, aussuchen könntest, welches würdest Du wählen?
9. Wenn Du Dir eine Schallplatte mit Musik aus einem bestimmten Land auswählen könntest, welche Schallplatte würdest Du wählen? Mit Musik aus ...
10. Wenn Du wählen könntest, welche europäische Sprache würdest Du lernen?
11. Wenn Du einen Brieffreund, dem Du ab und zu einen Brief schreibst, wählen könntest, aus welchem Land würdest Du ihn aussuchen?
12. Jedes Land hat seine Entstehungsgeschichte; von welchem Land interessiert diese Dich am meisten?
13. Wenn Du Briefmarken zur Auswahl hättest, die Briefmarken aus welchem Land würden Dich am meisten interessieren?
14. Wenn Du Dir einen Kalender mit Photographien aus verschiedenen Ländern aussuchen könntest, welchen Kalender würdest Du in Deinem Zimmer gerne aufhängen? Einen Kalender mit Photographien aus ...

Die Vpn wurden anhand dieser Fragen dazu aufgefordert, sich für mindestens eines der 12 Länder, die neben Deutschland jetzt schon zur EG gehören bzw. diese Zugehörigkeit anstreben (bzw. für eine ihrer Sprachen) zu entscheiden:
Belgien, Dänemark, England, Frankreich, Griechenland, Holland, Irland, Italien, Luxemburg, Portugal, Spanien, Türkei.

Der Fragebogen wurde bisher insgesamt 52 Vpn vorgelegt (30 PL, 22 N-PL), alle Hauptschüler der 8. Klasse, die alle 4 Jahre am Englischunterricht teilgenommen hatten. Die Plansprachschüler hatten zudem bis zu 2 Jahre an einer plansprachlichen Arbeitsgemeinschaft teilgenommen.

Die Schüler sollten entweder eine der Möglichkeiten ankreuzen oder durch die Ziffern 1,2,3 die drei bevorzugten Möglichkeiten in eine Prioritätsfolge bringen.

3. *Auswertung und Interpretation der Ergebnisse*

Bei der Auswertung wurden (implizit) von jeder Vp bei jeder Frage 6 Punkte verteilt. Diese entfielen bei Ankreuzung nur einer Wahlmöglichkeit auf diese gewählte Möglichkeit, bei mehreren wurden sie auf alle verteilt, wobei bei 3 Ankreuzungen die erste mit 3 Punkten, die zweite mit 2 und die dritte mit einem Punkt bewertet wurde. Wir erhielten aus den Antworten zu den Fragen 1–9 und 11–14 13 Felder mit je 12 Elementen (den Ländern) und aufgrund der Frage 10 ein Feld mit 10 Elementen (den Sprachen). Die Häufigkeitsverteilungen sind die jeweilige Verteilung der einzelnen Punkte. Wir berechneten für die einzelnen Felder die mittlere Information nach Gleichung

$$H = \operatorname{ld} N \quad \frac{\sum_i N_i \operatorname{ld} N_i}{N} + \frac{df}{1{,}3863 \cdot N}$$

(vgl. Frank, 1969, Bd. 2, S. 224). Dabei bedeutet N die Gesamtzahl von verteilten Punkten, N_i die Gesamtzahl der auf ein Land verteilten Punkte und df die Anzahl der Freiheitsgrade.

Der letzte Summand ist ein Korrekturglied, auf das Schmitt (1964) in seiner Dissertation hinweist; es dient zur genaueren Errechnung von mittleren Informationswerten H, wenn in Feldern statt der Wahrscheinlichkeiten nur die relativen Häufigkeiten bekannt sind. Da durch eine Häufung von Punkten bei wenigen Elementen die mittlere Information verringert wird, ist nach unserer Arbeitshypothese zu erwarten, daß die Werte von H bei der N-PL-Gruppe deutlich geringer sind als bei der PL-Gruppe. Die tabellarische Gegenüberstellung der errechneten Werte in der Tabelle bestätigt in 9 Fällen diese Erwartung;

Fragen-Nr.	$H_{N\text{-}PL}$/bit		H_{PL}/bit
1	3,08	$<$	3,20
2	3,03	$<$	3,24
3	3,01	$>$	2,78
4	3,32	$>$	3,27
5	2,07	$<$	2,88
6	2,53	$<$	3,12
7	3,07	$<$	3,18
8	3,29	$>$	3,12
9	1,84	$>$	1,11
10	2,37	$>$	2,06
11	2,85	$<$	3,14
12	2,70	$<$	3,3
13	2,92	$<$	3,19
14	2,66	$<$	3,04

dagegen traten bei 5 Fällen (Fragen 3, 4, 8, 9, 10) Ergebnisse auf, die von unserer theoretischen Annahme abweichen. Dafür sehen wir zunächst folgende z.T. meßtechnische, z.T. theoretische Erklärungsmöglichkeiten:

1. Die bisher geringe Anzahl von Vpn macht statistische Zufallsabweichungen nicht genügend unwahrscheinlich.
2. Die getroffene Auswahl der Fragen und damit die hinter ihnen stehenden Merkmale sind nicht repräsentativ für die Beschreibung der Ursachen der Interessen gegenüber Ländern.
3. Die weitergehende und gleichmäßigere Informierung der einen Versuchsgruppe (PL) führt bei manchen Merkmalen der Länder nicht zu einer gleichmäßigen Verteilung der effektiv getroffenen Entscheidungen in der Versuchsgruppe (PL), da u.U. die Frage auf schon bestehende Interessen stößt, aufgrund derer manche Entscheidungsmöglichkeiten von vorneherein ausgeschlossen und andere besonders nahegelegt werden. Man müßte in dem Fall nicht allgemein von einer (neutralen) Informierung sprechen, sondern von „Negativ-Information" bzw. „Positiv-Information", die beim Prozeß der Informationsaufnahme verarbeitet wird und die Entscheidungen steuert.

4. *Schlußfolgerung*

Fordert man – wie z.B. Frank (1962) – ein Staatswesen, „in welchem jeder beliebige selbständig denkende Staatsbürger ein Maximum an Freiheit genießt, ..." und sucht man auf dieser Grundlage nach Richtlinien für eine Europa-orientierte Erziehung und Bildung („Europädagogik") und akzeptiert man als empirisches Kontrollverfahren des Ausmaßes an Freiheit die Vorgehensweise von D. und A. Gabor (mit den eingangs genannten Einschränkungen), dann deuten schon die mitgeteilten ersten Ergebnisse der noch laufenden Untersuchung darauf hin, daß durch den zusätzlichen Plansprachunterricht anstelle einer Beschränkung auf den Englischunterricht beim Schüler eine zu begrüßende Freiheitsvergrößerung durch europäische „Horizonterweiterung" eintritt. Es bleibt zu untersuchen, ob dieser europädagogische Vorteil des Plansprachunterrichts für eine freiheitliche europäische Erziehung sich auch schon beim Sprachorientierungsunterricht in der Grundschule nachweisen läßt.

Schrifttum

Frank, H.: Über die wissenschaftliche und ideologische Komponente im Maß der Freiheit. Z. phil. Forschung, 1962, 16, 99–117
Frank, H.: Kybernetische Grundlagen der Pädagogik. Agis: Baden-Baden, 2. Auflg. 1969
Gabor, D.&A.: An Essay on the mathematical Theory of Freedom. Journ. Roy. Stat. Soc. Ser. A, 117, 1954, p. 31–59
Hofstätter, P. R.: Einführung in die Sozialpsychologie. Kröner: Stuttgart, 1966
Schmitt, E.: Der Binärprediktor. Diss. Karlsruhe, 1964
Thurstone, L. L.: The prediction of choice. Psychometrike, 4, 1945

Eingegangen am 15. August 1978

Anschrift der Verfasserin: Dr. Brigitte S. Meder, Heiersmauer 71, D-4790 Paderborn

Richtlinien für die Manuskriptabfassung

Es wird zur Beschleunigung der Publikation gebeten, Beiträge an die Schriftleitung in doppelter Ausfertigung einzureichen. Etwaige Tuschzeichnungen oder Photos brauchen nur einfach eingereicht zu werden.

Artikel von mehr als 12 Druckseiten Umfang können in der Regel nicht angenommen werden. Unverlangte Manuskripte können nur zurückgesandt werden, wenn Rückporto beiliegt. Es wird gebeten, für die Aufnahme in die internationale Knapptextbeilage „Homo kaj Informo" eine knappe, aber die wichtigsten neuen Ergebnisse des Beitrags für Fachleute verständlich wiedergebende Zusammenfassung (Umfang maximal 200 Wörter) in Internationaler, notfalls deutscher Sprache beizufügen.

Die verwendete Literatur ist, nach Autorennamen alphabetisch (verschiedene Werke desselben Autors chronologisch) geordnet, in einem Schrifttumsverzeichnis am Schluß des Beitrags zusammenzustellen. Die Vornamen der Autoren sind mindestens abgekürzt zu nennen. Bei selbständigen Veröffentlichungen sind Titel, Erscheinungsort und -jahr, womöglich auch Verlag, anzugeben. Zeitschriftenbeiträge werden vermerkt durch Name der Zeitschrift, Band, Seite (z. B. S. 317–324) und Jahr, in dieser Reihenfolge. (Titel der Arbeit soll angeführt werden.) Im selben Jahr erschienene Arbeiten desselben Autors werden durch den Zusatz „a", „b" etc. ausgezeichnet. Im Text soll grundsätzlich durch Nennung des Autorennamens und des Erscheinungsjahrs des zitierten Werkes (evtl. mit dem Zusatz „a" etc.), in der Regel aber nicht durch Anführung des ganzen Buchtitels zitiert werden. Wo es sinnvoll ist, sollte bei selbständigen Veröffentlichungen und längeren Zeitschriftenartikeln auch Seitenzahl oder Paragraph genannt werden. Anmerkungen sind zu vermeiden. Im übrigen wird auf die „Mindestgütekriterien für kybernetisch-pädagogische Originalarbeiten in deutscher Sprache" (abgedruckt u. a. in „Kybernetik und Bildung I", Verlagsgemeinschaft Schroedel/Schöningh, Hannover und Paderborn 1975) verwiesen, die von Schriftleitung und Herausgebern der Beurteilung der eingereichten Manuskripte sinngemäß zugrundegelegt werden.